

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM proteoin - proteoin search, using sw model

Run on: October 24, 2005, 13:10:07 ; Search time 170 Seconds
(without alignments)
659,678 Million cell updates/sec

Title: US-10-733-288b-4
Perfect score: 1142
Sequence: 1 IVGGRRARHPAFWVSLQL.....PDAPFAVQAQFVWVDSIIQR 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03 : *
1: uniprot_sprot: *
2: uniprot_tramb1: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1142	100.0	267	1	ELNE_HUMAN
2	905	79.2	282	1	Q8MJD1
3	873.5	76.5	285	2	Q61515
4	686	60.1	258	2	Q9GME1
5	659	57.7	207	2	Q9Z284
6	657	57.5	145	2	Q6LDP5
7	648.5	56.8	254	2	Q8K597
8	647.5	56.7	256	1	PRN3_HUMAN
9	646.5	56.6	237	1	Q6LBR2
10	644.5	56.4	254	1	PRN3_MOUSE
11	584.5	51.2	245	1	Q6DF10
12	489	42.8	251	1	CAP7_HUMAN
13	437	38.3	219	1	CAP7_PIG
14	392	34.3	264	2	Q8QGF6
15	392	34.3	264	2	Q8GPF5
16	387	33.9	258	2	Q867B0
17	384	33.6	265	2	Q66KX6
18	381	33.4	283	2	Q6UWY2
19	379	33.2	265	2	Q7SYX8
20	379	33.2	266	1	EL1_PIG
21	370.5	32.4	268	2	Q9W7Q2
22	370	32.4	258	2	Q61SW6
23	369	32.3	258	1	EL1_HUMAN
24	366	32.0	266	2	Q91X79
25	365	32.0	266	1	EL1_RAT
26	364	31.9	265	2	Q6NGC0
27	364	31.9	266	2	Q9D936
28	360.5	31.6	249	2	Q9W7Q1
29	358	31.3	245	1	MCT1_SHEEP
30	358	31.3	266	2	O46644
31	357	31.3	266	1	EL1_BOVIN

32	355	31.1	226	2	Q86SR2	Q86SR2 homo sapien
33	355	31.1	251	2	Q8GJN2	Q8GJN2 bos taurus
34	354	31.0	278	2	Q68FN6	Q68FN6 brachydanio
35	353.5	31.0	261	1	CATE_MOUSE	P28293 mus musculu
36	351.5	30.8	262	1	GRAA_HUMAN	P12544 homo sapien
37	351.5	30.8	266	2	Q6AZC0	Q6AZC0 brachydanio
38	350.5	30.7	267	2	Q6DCW4	Q6DCW4 brachydanio
39	349.5	30.6	226	1	DDN1_BOVIN	P80219 bos taurus
40	349	30.6	260	2	Q7S2C3	Q7S2C3 gallus gall
41	348.5	30.5	236	2	Q7S1G3	Q7S1G3 salmo salar
42	346.5	30.3	255	1	CATE_HUMAN	P08311 homo sapien
43	342	29.9	269	2	Q6AZF9	Q6AZF9 xenopus lae
44	340.5	29.8	257	1	GRAM_HUMAN	P51124 homo sapien
45	339.5	29.7	260	1	GRAA_MOUSE	P11032 mus musculu

ALIGNMENTS

RESULT 1	ELNE_HUMAN	STANDARD;	PRT;	267 AA.
AC	P08246; P09649;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Leukocyte elastase precursor (EC 3.4.21.37) (Neutrophil elastase) (PMN elastase) (Bone marrow serine protease) (Medullasin).			
GN	Name=ELA2;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=89374820; PubMed=2775493;			
RT	Farley D., Travis J., Salvesen G.;			
RT	"the human neutrophil elastase gene. Analysis of the nucleotide sequence reveals three distinct classes of repetitive DNA.";			
RL	Biol. Chem. Hoppe-Seyler 370:737-744 (1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=8806782; PubMed=3479752;			
RT	Nakanura H., Okano K., Aoki Y., Shimizu H., Naruto M.;			
RT	"Nucleotide sequence of human bone marrow serine protease (medullasin) gene.";			
RT	Nucleic Acids Res. 15:9601-9601 (1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=89008342; PubMed=2902087;			
RT	Takahashi H., Nuklwa T., Yoshimura K., Quick C.D., States D.J.,			
RT	Holmes M.D., Whang-Peng J., Knutsen T., Crystal R.G.;			
RT	"Structure of the human neutrophil elastase gene.";			
RL	J. Biol. Chem. 263:14739-14747 (1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=90211319; PubMed=2322278;			
RT	Okano K., Aoki Y., Shimizu H., Naruto M.;			
RT	"Functional expression of human leukocyte elastase (HLE)/medullasin in eukaryotic cells.";			
RL	Biochem. Biophys. Res. Commun. 167:1326-1332 (1990).			
RN	[5]			
RP	SEQUENCE FROM N.A. AND VARIANTS ILE-219; LEU-257 AND LEU-262.			
RA	Livingston R.J., Rieder M.J., Chung M.-W., Ritchie T.K., Olson A.N.,			
RA	Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S.,			
RA	Sherwood J.K., Sherwood A.M., Leithausen B.J., Nickerson D.A.;			
RT	"NIHES-SNPs, environmental genome project, NIHES ES15478, Department of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";			
RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE OF 30-267 FROM N.A.			
RA	MEDLINE=88032918; PubMed=2822677;			
RT	Okano K., Aoki Y., Sakurai T., Kajitani M., Kanai S., Shimazu T.,			

RA Shimizu H., Naruto M.;
 RT "Molecular cloning of complementary DNA for human medullasin: an
 RT inflammatory serine protease in bone marrow cells."; *Int J. Biochem.* 102:13-16(1987).
 RN [17]
 RN SEQUENCE OF 75-267 FROM N.A.
 RA MEDLINE=88115408; PubMed=3422232;
 RA Takahashi H., Nukima T., Baesser P.;
 RT "Myelomonocytic cell lineage expression of the neutrophil elastase
 RT gene."; *J. Biol. Chem.* 263:2543-2547(1988).
 RN [18]
 RN SEQUENCE OF 30-247.
 RA MEDLINE=87175647; PubMed=3550808;
 RA Sinha S., Matorek W., Kerr S., Giles J., Bode W., Travis J.;
 RT "Primary structure of human neutrophil elastase."; *Proc. Natl. Acad. Sci. U.S.A.* 84:2228-2232(1987).
 RN [19]
 RN SEQUENCE OF 262-267.
 RA MEDLINE=91315473; PubMed=1859409;
 RA Aoki Y., Hase T.;
 RT "The primary structure and elastolytic activity of medullasin (a
 RT serine protease of bone marrow)."; *Biochem. Biophys. Res. Commun.* 178:501-506(1991).
 RN [110]
 RN PRELIMINARY SEQUENCE OF 30-103.
 RA Travis J., Giles P.J., Porcelli L., Reilly C.F., Baugh R., Powers J.;
 RT (in) Protein degradation in health and disease, Ciba Foundation
 RL Symposium, pp.75:51-68, Excerpta Medica, Amsterdam and Oxford (1980).
 RN [111]
 RN SEQUENCE OF 30-49.
 RA MEDLINE=89315847; PubMed=2501794;
 RA Gabay J.E., Scott R.W., Campanelli D., Griffith J., Wilde C.,
 RA Marra M.N., Seeger M., Nathan C.F.;
 RT "Antibiotic proteins of human polymorphonuclear leukocytes."; *Proc. Natl. Acad. Sci. U.S.A.* 86:5610-5614(1989).
 RN [112]
 RN X-RAY CRYSTALLOGRAPHY (1.84 ANGSTROMS).
 RA MEDLINE=89098932; PubMed=2911584;
 RA Navia J.A., McKeever B.M., Springer J.P., Lin T.-Y., Williams H.R.,
 RA Fluder E.M., Dorn C.P., Hoogsteen K.;
 RT "Structure of human neutrophil elastase in complex with a peptide
 RT chromomethyl ketone inhibitor at 1.84-A resolution."; *Proc. Natl. Acad. Sci. U.S.A.* 86:7-11(1989).
 RN [113]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RA MEDLINE=86871660; PubMed=3391280; DOI=10.1016/0014-5793(88)80118-2;
 RA Wei A.-Z., Mayr I., Bode W.;
 RT "The refined 2.3-A crystal structure of human leukocyte elastase in a
 RT complex with a valine chloromethyl ketone inhibitor."; *FEBS Lett.* 234:367-373(1988).
 RN [114]
 RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RA MEDLINE=87053808; PubMed=3640709;
 RA Bode W., Wei A.-Z., Huber R., Meyer E., Travis J., Neumann S.;
 RT "X-ray crystal structure of the complex of human leukocyte elastase
 RT (PMN elastase) and the third domain of the turkey ovomucoid
 RT inhibitor."; *EMBO J.* 5:2453-2458(1986).
 RN [115]
 RN VARIANTS CH VAL-32; PHE-177 AND GLN-191.
 RA MEDLINE=20047772; PubMed=10581030; DOI=10.1038/70544;
 RA Horwitz M., Benson K.F., Person R.E., Aprikyan A.G., Dale D.C.;
 RT "Mutations in ELA2, encoding neutrophil elastase, define a 21-day
 RT biological clock in cyclic haematopoiesis."; *Nat. Genet.* 23:433-436(1999).
 CC -1- FUNCTION: Medullasin modifies the functions of natural killer
 CC cells, monocytes and granulocytes.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins, including elastin.
 CC Preferential cleavage: Val-I-Xaa > Ala-I-Xaa.
 CC -1- TISSUE SPECIFICITY: Bone marrow cells.
 CC -1- DISORD: Defects in ELA2 are a cause of cyclic haematopoiesis (CH)
 CC [MIM:162800]; also known as cyclic neutropenia. CH is an autosomal

CC dominant disease in which blood-cell production from the bone
 CC marrow oscillates with 21-day periodicity. Circulating neutrophils
 CC vary between almost normal numbers and zero. During intervals of
 CC neutropenia, affected individuals are at risk for opportunistic
 CC infection. Monocytes, platelets, lymphocytes and reticulocytes
 CC also cycle with the same frequency.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Elastase
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, J03545; AAA52378.1; -;
 CC EMBL, Y00477; CAA68537.1; -;
 CC EMBL, X05875; CAA29299.1; -;
 CC EMBL, X05875; CAA29300.1; ALT_INIT.
 CC EMBL, M20203; AAA36359.1; -;
 CC EMBL, M20199; AAA36359.1; JOINED.
 CC EMBL, M20200; AAA36359.1; JOINED.
 CC EMBL, M20201; AAA36359.1; JOINED.
 CC EMBL, AY596461; AAS89303.1; -;
 CC EMBL, M34379; AAA36173.1; -;
 CC EMBL, D00187; BAA00128.1; -;
 CC PIR, A31976; ELHDL.
 CC PDB, 1BOF; X-ray; A=30-247.
 CC PDB, 1H1B; X-ray; A/B=30-247.
 CC PDB, 1HNE; X-ray; E=30-247.
 CC PDB, 1PPF; X-ray; E=30-247.
 CC PDB, 1PRG; X-ray; E=30-247.
 CC MEROPS: S01.131, -;
 CC GeneW, HGNC:3509; ELA2.
 CC MIM, 130130; -;
 CC GO, GO:0004234; F:macrophage elastase activity; TAS.
 CC InterPro, IPR009003; Pept_Ser_Cys.
 CC InterPro, IPR001254; Peptidase_S1.
 CC InterPro, IPR00114; Peptidase_S1A.
 CC Pfam, PF00089; Trypsin, 1.
 CC PRINTS, PR00722; CHYMOTRYPSIN.
 CC PROSITE, PS0240; TRYPSIN_HIS, 1.
 CC PROSITE, PS00134; TRYPSIN_HIS, 1.
 CC PROSITE, PS00135; TRYPSIN_SER, 1.
 CC 3D-structure: Direct protein sequencing; Disease mutation;
 CC Glycoprotein; Hydrolyase; Polymorphism; Serine protease; Signal.
 CC FT SIGNAL 1 27
 CC FT PROPEP 28 29
 CC FT CHAIN 30 267
 CC FT ACT_SITE 30 70
 CC FT ACT_SITE 117 117
 CC FT ACT_SITE 202 202
 CC FT DISULFID 55 71
 CC FT DISULFID 151 208
 CC FT DISULFID 181 187
 CC FT DISULFID 198 223
 CC FT CARBOHYD 88 88
 CC FT CARBOHYD 124 124
 CC FT CARBOHYD 173 173
 CC FT VARIANT 32 32
 CC FT VARIANT 177 177
 CC FT VARIANT 191 191
 CC FT VARIANT 219 219
 CC FT VARIANT 257 257
 CC FT VARIANT 262 262
 CC
 CC N-linked (GlcNAc...) (Potential).
 CC N-linked (GlcNAc...)
 CC N-linked (GlcNAc...)
 CC G->V (in CH).
 CC /FTID=VAR_009538.
 CC V->F (in CH).
 CC /FTID=VAR_009539.
 CC R->Q (in CH).
 CC /FTID=VAR_009540.
 CC V->I.
 CC /FTID=VAR_019237.
 CC P->L.
 CC /FTID=VAR_019238.
 CC P->L.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2005, 13:01:06 : Search time 170 Seconds

(without alignments)
498.238 Million cell updates/sec

Title: US-10-733-288B-4

Perfect score: 1142
Sequence: 1 IGVGRRRAPPAWPFMWLSLQ.....PDAPAPVAGVWNIDSLIQR 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1142	100.0	238	7	ADJ68391 Human hea
2	1142	100.0	247	4	ABBS5525 Human ela
3	1142	100.0	267	1	AAPE8035 Sequence
4	1142	100.0	267	2	AAW64262 Human neu
5	1142	100.0	267	8	ADK42607 Human neu
6	1132	99.1	218	7	ADJ68281 Human hea
7	1123	98.3	218	5	AAU75897 Human leu
8	1073	94.0	257	5	ABP69417 Human pol
9	880.5	77.1	265	2	AAW70534 Human neu
10	877.5	76.8	265	2	AAW70534 Human neu
11	848.5	56.8	256	2	AAW82639 Mouse ser
12	848.5	56.8	256	2	AAW82639 MY17 prep
13	847.5	56.7	229	7	ABW02656 Human pre
14	847.5	56.7	229	7	ABW02654 Human mat
15	847.5	56.7	256	2	AAW45403 Deduced s
16	847.5	56.7	256	2	ABW02646 Human pre
17	846.5	56.6	229	7	ADW67204 Human hom
18	844.5	56.4	229	7	ABW02678 Human mat
19	844.5	56.4	253	8	ADM67203 Human mat
20	843.5	56.3	229	7	ABW02676 Human mat
21	842.5	56.3	229	7	ABW02677 Human mat
22	841.5	56.2	229	7	ABW02680 Human mat
23	841.5	56.2	237	2	AAW20509 Human pro
24	841.5	56.2	256	5	ABW06770 Human pro
25	841.5	56.2	256	8	ADQ14416 Human mye

26	641.5	56.2	256	8	ADR41702	ADR41702	Wegener's
27	578.5	50.7	215	2	AAW13036	AAW13036	Myeloplas
28	494.5	43.3	255	8	ADT16330	ADT16330	Human pro
29	494	43.3	222	2	AAW73210	AAW73210	CAP37 pro
30	492	43.1	225	7	ADE11551	ADE11551	Human mat
31	490	42.9	225	3	AAW71881	AAW71881	Human hep
32	489	42.8	222	2	AAW41935	AAW41935	Recombina
33	489	42.8	222	7	ADG42088	ADG42088	Human pol
34	489	42.8	225	2	AAW88362	AAW88362	Human mat
35	489	42.8	225	2	AAW21551	AAW21551	Human hep
36	489	42.8	225	2	AAW68118	AAW68118	Human hu
37	489	42.8	225	3	AAW71883	AAW71883	Human hep
38	489	42.8	225	3	AAW71876	AAW71876	Human mat
39	489	42.8	225	4	AAW71891	AAW71891	Human mat
40	489	42.8	225	7	ADE11587	ADE11587	Human mat
41	489	42.8	225	8	ADK42019	ADK42019	Human hep
42	489	42.8	232	2	AAW68364	AAW68364	Human pro
43	489	42.8	232	2	AAW21550	AAW21550	Human hep
44	489	42.8	232	2	AAW68120	AAW68120	Human hep
45	489	42.8	232	3	AAW71877	AAW71877	Human pre

ALIGNMENTS

RESULT 1	ADJ68391	standard; protein; 238 AA.
ID	ADJ68391	
XX	ADJ68391;	
AC		
XX		
DT	06-MAY-2004	(first entry)
XX		
DS	Human heat mitochondrial protein as a therapeutic target SegID197.	
XX		
KW	mitochondrial; human; screening assay; diabetes mellitus;	
KW	Huntington's disease; osteoarthritis;	
KW	Leber's hereditary optic neuropathy; LHON;	
KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;	
KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;	
KW	neuroprotective; nocturnal; antidiabetic; anticonvulsant; antithyroidic;	
KW	osteoplastic; ophthalmological; cytostatic.	
XX		
OS	Homo sapiens.	
XX		
FN	WO2003087768-A2.	
XX		
PD	23-OCT-2003.	
XX		
PF	04-APR-2003; 2003WC-US010870.	
XX		
PR	12-APR-2002; 2002US-0372843P.	
PR	17-JUN-2002; 2002US-0389987P.	
PR	20-SEP-2002; 2002US-0412418P.	
XX		
PA	(MITO-) MITOKOR.	
PA	(BUCK-) BUCK INST AGE RES.	
XX		
PI	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;	
PI	Warnock DE;	
XX		
DR	WPI; 2003-845369/78.	
XX		
PT	Identifying a mitochondrial target for drug screening assays and for	
PT	treating diseases associated with altered mitochondrial function,	
PT	comprises detecting a modified polypeptide in a sample and correlating	
PT	with the disease.	
PS	Claim 1; SEQ ID NO 197; 180pp: English.	
XX		
CC	This invention relates to novel mitochondrial targets that can be used	
CC	for therapeutic intervention in treating a disease associated with	
CC	altered mitochondrial function. Specifically, it refers to a method for	

CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteoprotic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

SQ Sequence 238 AA;

Query Match 100.0%; Score 1142; DB 7; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVGRRARPHAMPFWVSLQLRGHEFCATLLAPNFVMSAAHCVANVNVRAVVLGAHNL 60
DB 1 IVGRRARPHAMPFWVSLQLRGHEFCATLLAPNFVMSAAHCVANVNVRAVVLGAHNL 60
QY 61 SRREPTROVFAVQRIFFENGYPVNLNDIVIIQLNGSATINNAVQAOIPAGRRLGNGV 120
DB 61 SRREPTROVFAVQRIFFENGYPVNLNDIVIIQLNGSATINNAVQAOIPAGRRLGNGV 120
QY 121 QCLANGWGLLGRNRGIAVLOELNVTVTSLCRSNVCTLVGRGAGVCFGSGSPLVCN 180
DB 121 QCLANGWGLLGRNRGIAVLOELNVTVTSLCRSNVCTLVGRGAGVCFGSGSPLVCN 180
QY 181 GLIHGIASFVRGCGASGLYPDAFAPVPAQFVMWIDSIIQR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVPAQFVMWIDSIIQR 219

RESULT 2
AAB55525
ID AAB55525 standard; protein; 247 AA.
XX

AC AAB55525;

DT 07-MAR-2001 (first entry)

DE Human elastase variant protein SEQ ID NO:95.

XX Human; elastase; variant; substrate; mutant; mutagenesis; histidine;
KW human neutrophil elastase; H43A; cytostatic; proteolysis; ADAPT;
XX antibody-directed enzyme activated prodrug therapy.

OS Homo sapiens.

PN WO200066363-A2.

PD 16-NOV-2000.

PF 04-MAY-2000; 2000WO-US006692.

PR 05-MAY-1999; 99US-0132640P.

PA (GETH) GENENTECH INC.

PI Carter PJ, Dall'acqua W, Rodrigues M;

DR MPI: 2001-007389/01.

N-PSDB; AAC88022.

XX Elastase variant (H43A) having altered substrate specificity useful for
XX antibody-directed enzyme activated prodrug therapy.

PS Disclosure; Fig 2; 79pp; English.

XX The present invention describes a purified elastase variant (I) with an

CC amino acid sequence different from that of a precursor elastase, the
CC difference comprising a substitution of an active site histidine residue
CC corresponding to residue 43 in human neutrophil elastase with a different
CC amino acid residue so that (I) has substrate specificity substantially
CC different from the precursor elastase. (I) has cytostatic activity, and
CC can be used in antibody-directed enzyme activated prodrug therapy. The
CC elastase variant can be used to cleave a particular substrate, especially
CC those containing histidine residues at the substrate site. Site-specific
CC proteolysis is useful in therapeutic applications, e.g. for antibody-
CC directed enzyme activated prodrug therapy (ADAPT). AAC88022, AAC88023 and
CC AAB55432 to AAB55526 represent sequences used in the exemplification of
CC the present invention

SQ Sequence 247 AA;

Query Match 100.0%; Score 1142; DB 4; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.9e-67;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVGRRARPHAMPFWVSLQLRGHEFCATLLAPNFVMSAAHCVANVNVRAVVLGAHNL 60
DB 5 IVGRRARPHAMPFWVSLQLRGHEFCATLLAPNFVMSAAHCVANVNVRAVVLGAHNL 64
QY 61 SRREPTROVFAVQRIFFENGYPVNLNDIVIIQLNGSATINNAVQAOIPAGRRLGNGV 120
DB 65 SRREPTROVFAVQRIFFENGYPVNLNDIVIIQLNGSATINNAVQAOIPAGRRLGNGV 124
QY 121 QCLANGWGLLGRNRGIAVLOELNVTVTSLCRSNVCTLVGRGAGVCFGSGSPLVCN 180
DB 125 QCLANGWGLLGRNRGIAVLOELNVTVTSLCRSNVCTLVGRGAGVCFGSGSPLVCN 184
QY 181 GLIHGIASFVRGCGASGLYPDAFAPVPAQFVMWIDSIIQR 219
DB 185 GLIHGIASFVRGCGASGLYPDAFAPVPAQFVMWIDSIIQR 223

RESULT 3
AAP80335
ID AAP80335 standard; protein; 267 AA.
XX

AC AAP80335;

DT 25-MAR-2003 (revised)

DT 04-OCT-1990 (first entry)

XX Sequence of serine protease (SP) of human myeloid cell origin and leader
DE peptide.
XX Serine protease; myeloid cell; intravascular coagulation treatment;
KW enzyme.

OS Homo sapiens.

PN WO8806621-A.

PD 07-SEP-1988.

PF 26-FEB-1988; 88WO-JP000205.

PR 05-MAR-1987; 87JP-00050676.

PA (TORA) TORAY IND INC.

PI Aoki Y, Aoki Y.

DR MPI: 1988-271161/38.

N-PSDB; AAP80315.

XX Aoki Y, Okano K, Naruto M, Shimizu H, Nakamura H;
XX MPI: 1988-271161/38.

PS Disclosure; AAP80315.

XX The present invention describes a purified elastase variant (I) with an

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2005, 12:59:58 : Search time 42 Seconds
(without alignments)
389,242 Million cell updates/sec

Title: US-10-733-288b-4

Perfect score: 1142
Sequence: 1 IVGGRRARRHAMPFWSLQL.....PDAPFVAQFVWKIDSIQR 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/aa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/aa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/aa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/aa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/aa/backfile1.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1142	100.0	238	3 US-08-944-483-31	Sequence 31, Appl
2	1142	100.0	238	6 5180819-4	Patent No. 5180819
3	1142	100.0	238	6 5180819-4	Patent No. 5180819
4	1142	100.0	257	2 US-08-978-404B-46	Sequence 46, Appl
5	792	69.4	271	3 US-09-578-303-6	Sequence 6, Appl
6	648.5	56.8	229	4 US-08-395-456C-27	Sequence 27, Appl
7	648.5	56.8	221	4 US-08-395-456C-25	Sequence 25, Appl
8	648.5	56.8	226	4 US-08-395-456C-23	Sequence 23, Appl
9	648.5	56.8	229	4 US-08-487-453A-23	Sequence 23, Appl
10	647.5	56.7	229	2 US-08-394-600B-20	Sequence 20, Appl
11	647.5	56.7	229	4 US-08-395-456C-20	Sequence 20, Appl
12	647.5	56.7	229	4 US-08-487-453A-20	Sequence 20, Appl
13	647.5	56.7	229	5 PCT-US95-02513-20	Sequence 20, Appl
14	647.5	56.7	226	2 US-08-230-428B-4	Sequence 4, Appl
15	641.5	56.2	229	3 US-08-944-483-30	Sequence 30, Appl
16	641.5	56.2	256	4 US-09-949-016-6271	Sequence 6271, Ap
17	590	51.7	214	6 5180819-3	Patent No. 5180819
18	590	51.7	214	6 5180819-3	Patent No. 5180819
19	581.5	50.9	215	6 5180819-2	Patent No. 5180819
20	581.5	50.9	215	6 5180819-2	Patent No. 5180819
21	494	43.3	222	2 US-08-491-204A-18	Sequence 18, Appl
22	489	42.8	232	1 US-07-969-931-9	Sequence 9, Appl
23	489	42.8	232	1 US-07-969-931-9	Sequence 9, Appl
24	489	42.8	225	3 US-08-944-483-32	Sequence 32, Appl
25	489	42.8	251	4 US-09-949-016-6112	Sequence 6112, Ap
26	489	42.8	255	4 US-09-949-016-9690	Sequence 9690, Ap
27	486	42.6	221	2 US-08-925-708-1	Sequence 1, Appl

28	433	37.9	219	2 US-08-925-708-2	Sequence 2, Appl
29	381	33.4	283	4 US-09-244-111-2	Sequence 2, Appl
30	377	33.0	240	1 US-08-278-091-11	Sequence 11, Appl
31	377	33.0	240	1 US-08-483-859-11	Sequence 11, Appl
32	377	33.0	240	1 US-08-472-173-11	Sequence 11, Appl
33	377	33.0	240	2 US-08-487-167-11	Sequence 11, Appl
34	377	33.0	240	2 US-08-482-816-11	Sequence 11, Appl
35	377	33.0	240	2 US-08-296-149-11	Sequence 11, Appl
36	377	33.0	240	2 US-08-801-499-11	Sequence 11, Appl
37	377	33.0	240	2 US-08-615-271-11	Sequence 11, Appl
38	377	33.0	240	3 US-09-074-660-11	Sequence 11, Appl
39	377	33.0	240	3 US-09-074-659-11	Sequence 11, Appl
40	377	33.0	240	3 US-09-106-468-11	Sequence 11, Appl
41	377	33.0	240	3 US-09-106-468-11	Sequence 11, Appl
42	377	33.0	240	3 US-09-106-467-11	Sequence 11, Appl
43	364	31.9	133	4 US-09-949-016-7471	Sequence 7471, Ap
44	351.5	30.8	224	3 US-08-944-483-46	Sequence 46, Appl
45	351.5	30.8	268	4 US-09-949-016-10712	Sequence 10712, A

ALIGNMENTS

RESULT 1
US-08-944-483-31
; Sequence 31, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLAS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: No. 62324566
US-08-944-483-31

Query Match 100.0%; Score 1142; DB 3; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.7e-123;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
DB 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
OY 61 SRRPTROYFAVQRIFFENGYPDVLNDIVIIQLNGSATINNAVQALPAQGRRLGNGV 120
DB 61 SRRPTROYFAVQRIFFENGYPDVLNDIVIIQLNGSATINNAVQALPAQGRRLGNGV 120
OY 121 QCLANGWGLGRRNGIASVLOELNVTVTSLCRRSNVCTLVGRQAGVCFGDSGPLVCN 180
DB 121 QCLANGWGLGRRNGIASVLOELNVTVTSLCRRSNVCTLVGRQAGVCFGDSGPLVCN 180
OY 181 GLIHGIASFVRGCGASGLYPDAFAPVAQFVNMIDSIQR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVAQFVNMIDSIQR 219

RESULT 2

5180819-4
PATENT No. 5180819
APPLICANT: CAYRE, YVON
TITLE OF INVENTION: PURIFIED MYELOBLASTIN, NUCLEIC ACID
MOLECULE ENCODING SAME, AND USES THEREOF
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/455,614
FILING DATE: 22-DEC-1989
SEQ ID NO:4
LENGTH: 238
5180819-4

Query Match 100.0%; Score 1142; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.7e-123;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
DB 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
OY 61 SRRPTROYFAVQRIFFENGYPDVLNDIVIIQLNGSATINNAVQALPAQGRRLGNGV 120
DB 61 SRRPTROYFAVQRIFFENGYPDVLNDIVIIQLNGSATINNAVQALPAQGRRLGNGV 120
OY 121 QCLANGWGLGRRNGIASVLOELNVTVTSLCRRSNVCTLVGRQAGVCFGDSGPLVCN 180
DB 121 QCLANGWGLGRRNGIASVLOELNVTVTSLCRRSNVCTLVGRQAGVCFGDSGPLVCN 180
OY 181 GLIHGIASFVRGCGASGLYPDAFAPVAQFVNMIDSIQR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVAQFVNMIDSIQR 219

RESULT 3

5180819-4
PATENT No. 5180819
APPLICANT: CAYRE, YVON
TITLE OF INVENTION: PURIFIED MYELOBLASTIN, NUCLEIC ACID
MOLECULE ENCODING SAME, AND USES THEREOF
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/455,614
FILING DATE: 22-DEC-1989
SEQ ID NO:4
LENGTH: 238
5180819-4

Query Match 100.0%; Score 1142; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.7e-123;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
DB 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
OY 61 SRRPTROYFAVQRIFFENGYPDVLNDIVIIQLNGSATINNAVQALPAQGRRLGNGV 120
DB 61 SRRPTROYFAVQRIFFENGYPDVLNDIVIIQLNGSATINNAVQALPAQGRRLGNGV 120
OY 121 QCLANGWGLGRRNGIASVLOELNVTVTSLCRRSNVCTLVGRQAGVCFGDSGPLVCN 180
DB 121 QCLANGWGLGRRNGIASVLOELNVTVTSLCRRSNVCTLVGRQAGVCFGDSGPLVCN 180
OY 181 GLIHGIASFVRGCGASGLYPDAFAPVAQFVNMIDSIQR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVAQFVNMIDSIQR 219

RESULT 4

US-08-978-404B-46
Sequence 46, Application US/08978404B

PATENT No. 5968782
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5968782e
US-08-978-404B-46

Query Match 100.0%; Score 1142; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 4.3e-123;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 60
DB 30 IVGRRARPHAMPFVMSIQLRGHFCGATLLIAPNFVMSAAHCVANNVRAVVLGAHNL 89

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2005, 12:59:58 ; Search time 164 Seconds
(without alignments)
557.524 Million cell updates/sec

Title: US-10-733-288b-4

Perfect score: 1142

Sequence: 1 IVGGRRARHPFWSLQ.....PDAPVAGVNMIDSIQR 219

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1862986 seqs, 417506248 residues

Total number of hits satisfying chosen parameters: 1862986

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10D_PUBCOMB.pep:*
17: /cgn2_6/prodata/1/pubppa/US10E_PUBCOMB.pep:*
18: /cgn2_6/prodata/1/pubppa/US10F_PUBCOMB.pep:*
19: /cgn2_6/prodata/1/pubppa/US11_PUBCOMB.pep:*
20: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB.pep:*
21: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
22: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1142	100.0	238	11	US-09-789-210-31
2	1142	100.0	238	16	US-10-408-765A-197
3	1142	100.0	240	18	US-10-970-794-8
4	1142	100.0	267	17	US-10-931-153-22
5	1142	100.0	267	18	US-10-970-794-7
6	1133	99.2	219	16	US-10-733-288A-4
7	1132	99.2	218	16	US-10-408-765A-87
8	1123	98.3	218	9	US-09-861-708-3
9	1114	97.5	218	16	US-10-733-288A-3
10	641.5	56.2	229	11	US-09-789-210-30
11	489	42.8	222	15	US-10-423-311-1

12	489	42.8	225	11	US-09-789-210-32	Sequence 32, Appl
13	489	42.8	225	17	US-10-868-577A-30	Sequence 30, Appl
14	489	42.8	225	17	US-10-868-577A-38	Sequence 38, Appl
15	425	37.2	199	14	US-10-384-474-8	Sequence 8, Appl
16	425	37.0	199	15	US-10-423-311-2	Sequence 2, Appl
17	385	33.7	284	9	US-09-888-615-112	Sequence 112, Appl
18	381	33.4	283	10	US-09-946-374-111	Sequence 111, Appl
19	381	33.4	283	13	US-10-052-568-272	Sequence 272, Appl
20	381	33.4	283	14	US-10-174-599-272	Sequence 272, Appl
21	381	33.4	283	14	US-10-176-788-272	Sequence 272, Appl
22	381	33.4	283	14	US-10-175-737-272	Sequence 272, Appl
23	381	33.4	283	14	US-10-174-581-272	Sequence 272, Appl
24	381	33.4	283	14	US-10-176-483-272	Sequence 272, Appl
25	381	33.4	283	14	US-10-176-749-272	Sequence 272, Appl
26	381	33.4	283	14	US-10-176-914-272	Sequence 272, Appl
27	381	33.4	283	14	US-10-176-915-272	Sequence 272, Appl
28	381	33.4	283	14	US-10-173-706-272	Sequence 272, Appl
29	381	33.4	283	14	US-10-175-738-272	Sequence 272, Appl
30	381	33.4	283	14	US-10-175-752-272	Sequence 272, Appl
31	381	33.4	283	14	US-10-176-482-272	Sequence 272, Appl
32	381	33.4	283	14	US-10-176-757-272	Sequence 272, Appl
33	381	33.4	283	14	US-10-176-913-272	Sequence 272, Appl
34	381	33.4	283	14	US-10-180-552-272	Sequence 272, Appl
35	381	33.4	283	14	US-10-180-557-272	Sequence 272, Appl
36	381	33.4	283	14	US-10-173-700-272	Sequence 272, Appl
37	381	33.4	283	14	US-10-174-572-272	Sequence 272, Appl
38	381	33.4	283	14	US-10-174-579-272	Sequence 272, Appl
39	381	33.4	283	14	US-10-174-582-272	Sequence 272, Appl
40	381	33.4	283	14	US-10-174-588-272	Sequence 272, Appl
41	381	33.4	283	14	US-10-175-739-272	Sequence 272, Appl
42	381	33.4	283	14	US-10-175-740-272	Sequence 272, Appl
43	381	33.4	283	14	US-10-175-743-272	Sequence 272, Appl
44	381	33.4	283	14	US-10-176-488-272	Sequence 272, Appl
45	381	33.4	283	14	US-10-176-492-272	Sequence 272, Appl

ALIGNMENTS

RESULT 1
US-09-789-210-31
Sequence 31, Application US/097892210
Publication No. US20040241646A1

GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GRANADOS, EDWARD N.
KLASS, MICHAEL R.
RUSSELL, JOHN C.
STEWART, KENT D.
STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES OF THE PROSPATE

NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/789,210
FILING DATE: 20-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/944,483
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-789-210-31

Query Match 100.0%; Score 1142; DB 11; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.3e-110;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVGRRARPHAMPFVMSIQLRGSHFCGATLLAPNFVMSAAHCVANVNAVAHVVLGAHNL 60
DB 1 IVGRRARPHAMPFVMSIQLRGSHFCGATLLAPNFVMSAAHCVANVNAVAHVVLGAHNL 60
QY 61 SRREPTROVFAVORIFENGYPDNLNDIVILQLNGSATINNAVQAOQLPAQGRRLGNGV 120
DB 61 SRREPTROVFAVORIFENGYPDNLNDIVILQLNGSATINNAVQAOQLPAQGRRLGNGV 120
QY 121 OCLAMGWLGRNRGIAASVLOELNVTVTSLCRSNVCTLVGRGAGVCFGDSGSPLYCN 180
DB 121 OCLAMGWLGRNRGIAASVLOELNVTVTSLCRSNVCTLVGRGAGVCFGDSGSPLYCN 180
QY 181 GLIHGIASFVRGCGASGLYPDAFAPVAFVNMWIDSIIR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVAFVNMWIDSIIR 219

RESULT 2
US-10-408-765A-197
Sequence 197, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Rany, Bojin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Marnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ IDS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 197
LENGTH: 238
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-197

Query Match 100.0%; Score 1142; DB 16; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.3e-110;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVGRRARPHAMPFVMSIQLRGSHFCGATLLAPNFVMSAAHCVANVNAVAHVVLGAHNL 60
DB 1 IVGRRARPHAMPFVMSIQLRGSHFCGATLLAPNFVMSAAHCVANVNAVAHVVLGAHNL 60

QY 61 SRREPTROVFAVORIFENGYPDNLNDIVILQLNGSATINNAVQAOQLPAQGRRLGNGV 120
DB 61 SRREPTROVFAVORIFENGYPDNLNDIVILQLNGSATINNAVQAOQLPAQGRRLGNGV 120
QY 121 OCLAMGWLGRNRGIAASVLOELNVTVTSLCRSNVCTLVGRGAGVCFGDSGSPLYCN 180
DB 121 OCLAMGWLGRNRGIAASVLOELNVTVTSLCRSNVCTLVGRGAGVCFGDSGSPLYCN 180
QY 181 GLIHGIASFVRGCGASGLYPDAFAPVAFVNMWIDSIIR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVAFVNMWIDSIIR 219

RESULT 3
US-10-970-794-8
Sequence 8, Application US/10970794
Publication No. US20050136398A1
GENERAL INFORMATION:
APPLICANT: McDermott, Martin
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING THERAPEUTIC COMPOUNDS
FILE REFERENCE: 18477.043
CURRENT APPLICATION NUMBER: US/10/970,794
PRIOR FILING DATE: 2004-10-22
PRIOR APPLICATION NUMBER: 60/514,241
PRIOR FILING DATE: 2003-10-24
PRIOR APPLICATION NUMBER: 60/514,299
PRIOR FILING DATE: 2003-10-24
PRIOR APPLICATION NUMBER: 60/513,532
PRIOR FILING DATE: 2003-10-24
PRIOR APPLICATION NUMBER: 60/513,542
PRIOR FILING DATE: 2003-10-24
PRIOR APPLICATION NUMBER: 60/514,894
PRIOR FILING DATE: 2003-10-29
PRIOR APPLICATION NUMBER: 60/514,925
PRIOR FILING DATE: 2003-10-29
PRIOR APPLICATION NUMBER: To Be Assigned
PRIOR FILING DATE: 2004-10-22
NUMBER OF SEQ IDS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 240
TYPE: PRT
ORGANISM: Homo sapiens
US-10-970-794-8

Query Match 100.0%; Score 1142; DB 18; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.3e-110;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVGRRARPHAMPFVMSIQLRGSHFCGATLLAPNFVMSAAHCVANVNAVAHVVLGAHNL 60
DB 3 IVGRRARPHAMPFVMSIQLRGSHFCGATLLAPNFVMSAAHCVANVNAVAHVVLGAHNL 62
QY 61 SRREPTROVFAVORIFENGYPDNLNDIVILQLNGSATINNAVQAOQLPAQGRRLGNGV 120
DB 61 SRREPTROVFAVORIFENGYPDNLNDIVILQLNGSATINNAVQAOQLPAQGRRLGNGV 122
QY 121 OCLAMGWLGRNRGIAASVLOELNVTVTSLCRSNVCTLVGRGAGVCFGDSGSPLYCN 180
DB 121 OCLAMGWLGRNRGIAASVLOELNVTVTSLCRSNVCTLVGRGAGVCFGDSGSPLYCN 182
QY 181 GLIHGIASFVRGCGASGLYPDAFAPVAFVNMWIDSIIR 219
DB 181 GLIHGIASFVRGCGASGLYPDAFAPVAFVNMWIDSIIR 221

RESULT 4
US-10-931-153-22
Sequence 22, Application US/10931153
Publication No. US20050089515A1
GENERAL INFORMATION:
APPLICANT: Ley, Arthur C.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2005, 13:09:21 ; Search time 39 Seconds

(without alignments)
540.294 Million cell updates/sec

Title: US-10-733-288b-4

Perfect score: 1142
Sequence: 1 TVGRRARRHAMPFVSLQ.....PDARFVAQPVWMDISIQ 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR 79:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1142	100.0	267	1	ELHUL
2	873.5	76.5	265	2	leukocyte elastase
3	646.5	56.6	256	1	neutrophil elastase
4	489	42.8	251	1	proteinase 3 (EC 3
5	437	38.3	219	1	azurocidin precurs
6	379	32.2	256	1	azurocidin - pig
7	368	32.2	258	4	pancreatic elastase
8	368	32.2	267	4	pancreatic elastase
9	365	31.0	261	4	probable pancreatic
10	353.5	30.8	262	1	pancreatic elastase
11	351.5	30.6	262	1	cathepsin G (EC 3.
12	349.5	30.6	262	1	granzyme A (EC 3.4
13	346.5	30.3	255	2	cathepsin G (EC 3.
14	341	29.9	244	2	duodenase - bovine
15	339.5	29.7	257	2	cathepsin G (EC 3.
16	339.5	29.6	260	2	mast cell proteinase
17	337.5	29.6	249	2	granzyme A (EC 3.4
18	336.5	29.5	259	2	granzyme M (EC 3.4
19	335.5	29.4	253	1	pancreatic elastase
20	335	29.3	258	1	pancreatic elastase
21	333.5	29.2	259	2	complement factor
22	332.5	29.1	271	1	complement factor
23	332	29.1	236	1	pancreatic elastase
24	332	29.1	246	1	complement factor
25	331.5	29.0	271	2	complement factor
26	329.5	28.9	251	2	pancreatic elastase
27	329.5	28.9	265	2	pancreatic elastase
28	329.5	28.9	269	2	pancreatic elastase
29	329	28.8	244	2	chymase (EC 3.4.21

30	329	28.8	246	2	A36678	mast cell proteinase
31	328.5	28.8	226	1	KCFP	bradykinin (EC 3.4
32	322	28.2	258	2	A45161	serine proteinase
33	321	28.1	246	1	A46504	chymase (EC 3.4.21
34	321	28.1	810	1	PLHU	plasma (EC 3.4.21
35	320	28.0	268	2	S68825	pancreatic elastase
36	317	27.8	246	2	S64707	chymase (EC 3.4.21
37	316.5	27.7	271	2	S29339	chymotrypsin (EC 3
38	316.5	27.7	261	1	A61021	granzyme B (EC 3.4
39	315.5	27.6	812	1	PLBO	plasma (EC 3.4.21
40	312	27.3	247	2	S64708	chymase (EC 3.4.21
41	312	27.3	268	2	S68826	pancreatic elastase
42	312	27.3	455	2	A61545	plasma (EC 3.4.21
43	312	27.3	810	2	B30848	plasma (EC 3.4.21
44	311.5	27.3	810	2	I46260	plasma (EC 3.4.21
45	310.5	27.2	243	2	A35671	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

leukocyte elastase (EC 3.4.21.37) precursor [validated] - human
N:Alternate names: inflammatory serine proteinase; medullasin; neutrophil elastase
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1990 #sequence, revision 30-Jun-1990 #text, change 09-Jul-2004
C:Accession: A31976; S04954; S06241; A27064; S00631; A28370; A34570; A05293; A25907; S14
R:Takahashi, H.; Nukitwa, T.; Yoshimura, K.; Quick, C.D.; States, D.J.; Holmes, M.D.; War
J. Biol. Chem. 263, 14739-14747, 1988
A:Title: Structure of the human neutrophil elastase gene.
A:Reference number: A31976; PMID:89008342; PMID:2902087
A:Accession: A31976
A:Molecule type: DNA
A:Residues: 1-267 <TAK>
A:Cross-references: UNIPROT:P08246; GB:M20203; GB:J04032; NID:G189147; PIDN:AAA36359.1; I
R:Farley, D.; Travis, J.; Salvesen, G.
Biol. Chem. Hoppe-Seyler 370, 737-744, 1989
A:Title: The human neutrophil elastase gene. Analysis of the nucleotide sequence reveals
A:Reference number: S04954; PMID:89374820; PMID:2775493
A:Accession: S04954
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-267 <PAR>
R:Nakamura, H.; Okano, K.; Aoki, Y.; Shimizu, H.; Naruto, M.
Nucleic Acids Res. 15, 9601-9602, 1987
A:Title: Nucleotide sequence of human bone marrow serine protease (medullasin) gene.
A:Reference number: S06241; PMID:88067782; PMID:3479752
A:Accession: S06241
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-267 <NAK>
A:Cross-references: EMBL:Y00477; NID:G34529; PIDN:CAA68537.1; PID:G296665
R:Okano, K.; Aoki, Y.; Sakurai, T.; Kajitani, M.; Kanai, S.; Shimizu, T.; Shimizu, H.; Ne
J. Biochem. 102, 13-16, 1987
A:Title: Molecular cloning of complementary DNA for human medullasin: an inflammatory ser
A:Reference number: A27064; PMID:88022918; PMID:2822677
A:Accession: A27064
A:Molecule type: mRNA
A:Residues: 30-267 <OKA>
A:Cross-references: EMBL:X05875; NID:G34532; PIDN:CAA29300.1; PID:G1335212
R:Farley, D.; Salvesen, G.; Travis, J. 3-7, 1988
Biol. Chem. Hoppe-Seyler 369(suppl.), 3-7, 1988
A:Title: Molecular cloning of human neutrophil elastase.
A:Reference number: S00631; PMID:89076526; PMID:2462434
A:Accession: S00631
A:Molecule type: mRNA
A:Residues: 123-267 <FA2>
A:Cross-references: GB:M27783; NID:G182055; PIDN:AAA35792.1; PID:G182056
A:Note: the authors translated the codon TTC for residue 218 as Pro
R:Takahashi, H.; Nukitwa, T.; Bassett, P.; Crystal, R.G.
J. Biol. Chem. 263, 2543-2547, 1988

A>Title: Myelomonocytic cell lineage expression of the neutrophil elastase gene.
A:Reference number: A28370; MUID:88115408; PMID:3422232
A:Accession: A28370
A:Molecule type: mRNA
A:Residues: 75-267 <TA2>
A:Cross-references: GB:J03545; NID:9182050; PIDN:AAA52378.1; PID:9182051
B:Okano, K.; Aoki, Y.; Shimizu, H.; Naruto, M.
Biochem. Biophys. Res. Commun. 167, 1326-1332, 1990
A>Title: Functional expression of human leukocyte elastase (HLE)/medullasin in eukaryotic
A:Reference number: A35570; MUID:90211319; PMID:2322278
A:Accession: A34570
A:Molecule type: mRNA
A:Residues: 1-267 <OK2>
A:Cross-references: GB:M34379; NID:9187116; PIDN:AAA6173.1; PID:9307123
R:Travis, J.; Giles, P.J.; Porcellini, L.; Reilly, R.; Baugh, R.; Powers, J.
in Protein Degradation in Health and Disease, Ciba Foundation Symposium 75, 51-68, 1980
A:Reference number: A94428
A:Accession: A05293
A:Molecule type: protein
A:Residues: 30-66, 'G', 68-73, 'D', 75, 78-82, 'E', 84-89, 'T', 91-94, 'PT', 97-100, 'L', 102-103 <TR
A:Experimental source: neutrophil granulocytes
R:Simha, S.; Mewcock, W.; Karr, S.; Giles, J.; Bode, W.; Travis, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 2228-2232, 1987
A>Title: Primary structure of human neutrophil elastase.
A:Reference number: A25907; MUID:87175647; PMID:3550808
A:Accession: A25907
A:Molecule type: protein
A:Residues: 30-247 <SIN>
A>Note: the sequence from Fig. 1 is inconsistent with that from Fig. 2 in having 107-Asp
R:Green, B.G.; Weston, H.; Ashe, B.M.; Doherty, J.; Finke, P.; Hagmann, W.; Lark, M.; Ma
Arch. Biochem. Biophys. 266, 284-292, 1991
A>Title: PMN elastases: a comparison of the specificity of human isozymes and the enzyme
A:Reference number: S14736; MUID:91378304; PMID:1897955
A:Accession: S14736
A:Molecule type: protein
A:Residues: 30-50 <GRE>
C:Comment: This is a lysosomal proteinase found in the azurophil granules of neutrophils
C:Comment: This elastase cleaves preferentially bonds after Ala and Val. It is believed
C:Genetics:
A:Gene: GDB:ELA2
A:Cross-references: GDB:118792; OMIM:130130
A:Map position: 19p13.3-19p13.3
A:Introns: 23/1; 75/2; 122/3; 199/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: emphysema; glycoprotein; hydrolysis; leukocyte; lysosome; rheumatoid arthritis
F:1-27/Domain: signal sequence #status predicted <SIS>
F:28-29/Domain: propeptide #status predicted <PRO>
F:30-247/Product: leukocyte elastase #status experimental <MAT>
F:248-267/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:55-71, 151-208, 181-187, 198-223/Dissulfide bonds: #status experimental
F:70, 117/202/Active site: His, Asp, Ser #status predicted
F:88/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:124, 173/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 1142; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.1e-97;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IYGGRRARHPMPWVSLOLRGHEFCGATLIAPNFMMSAAHCYANNVRAVRVIGAHNL 60
DB 30 IYGGRRARHPMPWVSLOLRGHEFCGATLIAPNFMMSAAHCYANNVRAVRVIGAHNL 89
QY 61 SRREPTROVFAVORIFENGYPDVNLNDIVILQNGSATINNAVQALPAGGRRLGNGV 120
DB 90 SRREPTROVFAVORIFENGYPDVNLNDIVILQNGSATINNAVQALPAGGRRLGNGV 149
QY 121 QCLANGWGLGRNRGIAVLOELNVTVTSLCRSNVCTLVGRQAGVCFGDSGPLVCN 180
DB 150 QCLANGWGLGRNRGIAVLOELNVTVTSLCRSNVCTLVGRQAGVCFGDSGPLVCN 209
QY 181 GLIHGIAFVRGCGASGLYPDAPFAPVQFVNWIDSIOR 219

DB 210 GLIHGIAFVRGCGASGLYPDAPFAPVQFVNWIDSIOR 248
RESULT 2
148679
neutrophil elastase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #txc_change 09-Jul-2004
A:Accession: 148679
R:Nuchprayoon, I.; Meyers, S.; Scott, L.M.; Suzow, J.; Hiebert, S.; Friedman, A.D.
Mol. Cell. Biol. 14, 5558-5568, 1994
A>Title: FEBP2/CBF, the murine homolog of the human myeloid AML1 and PBFB2 beta/CBF beta
1s.
A:Reference number: 148679; MUID:94309676; PMID:8035830
A:Accession: 148679
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-265 <RES>
A:Cross-references: UNIPROT:Q61515; EMBL:U04962; NID:9452556; PIDN:AAB60670.1; PID:945255
C:Genetics:
A:Introns: 22/1; 74/2; 121/3; 199/3
C:Superfamily: trypsin; trypsin homology
F:29-242/Domain: trypsin homology <TRI>

Query Match 76.5%; Score 873.5; DB 2; Length 265;
Best Local Similarity 75.9%; Pred. No. 5.8e-73;
Matches 167; Conservative 20; Mismatches 33; Indels 1; Gaps 1;

QY 1 IYGGRRARHPMPWVSLOLRGHEFCGATLIAPNFMMSAAHCYANNVRAVRVIGAHNL 60
DB 29 IYGGRRARHPMPWVSLOLRGHEFCGATLIAPNFMMSAAHCYANNVRAVRVIGAHNL 88
QY 61 SRREPTROVFAVORIFENGYPDVNLNDIVILQNGSATINNAVQALPAGGRRLGNGV 120
DB 89 RROERTROTFSVQGFENGYPDSQULNDIVILQNGSATINNAVQALPAGGRRLGNGV 148
QY 121 QCLANGWGLGRNRGIAVLOELNVTVTSLC-RRSNVCTLVGRQAGVCFGDSGPLVC 179
DB 149 QCLANGWGLGRNRGIAVLOELNVTVTSLC-RRSNVCTLVGRQAGVCFGDSGPLVC 208
QY 180 NGLIHGIAFVRGCGASGLYPDAPFAPVQFVNWIDSIOR 219
DB 209 NNLVQIGDSFTRGGCGSLYPDAPFAPVQFVNWIDSIOR 248

RESULT 3
PRH03
Proteinase 3 (EC 3.4.21.-) precursor [validated] - human
N:Alternate names: AGP7; C-ANCA antigen; neutrophil proteinase 4; p29; Wegener's granulon
N:Contains: myeloblastin
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #txc_change 09-Jul-2004
A:Accession: A45080; B46268; A45983; JH0331; A33751; S11091; A61176; A60148; A43982; A43
R:Sturrock, A.B.; Franklin, K.F.; Rao, G.; Warshall, B.C.; Rebentisch, W.B.; Demons, R.S.
J. Biol. Chem. 267, 21193-21199, 1992
A>Title: Structure, chromosomal assignment, and expression of the gene for proteinase-3.
A:Reference number: A45080; MUID:93016043; PMID:1400430
A:Accession: A45080
A:Molecule type: DNA
A:Residues: 1-254, 'P', <STU>
A:Cross-references: UNIPROT:P24158; GB:M97911
A>Note: sequence extracted from NCBI backbone (NCBI:116205)
R:Zimmer, M.; Medcalf, R.L.; Fink, T.M.; Matzmann, C.; Lichter, P.; Jenne, D.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 8215-8219, 1992
A>Title: Three human elastase-like genes coordinately expressed in the myelomonocyte line
A:Reference number: A46268; MUID:92390417; PMID:1518649
A:Accession: B46268
A:Molecule type: DNA
A:Residues: 1-118, 'V', 120-134, 'AT', 137-256 <ZIM>
A>Note: sequence extracted from NCBI backbone (NCBI:112898, NCBI:112900, NCBI:112902,
R:Labbaye, C.; Musette, P.; Cayre, Y.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 9253-9256, 1991
A>Title: Wegener autoantigen and myeloblastin are encoded by a single mRNA.